ST 793 : Solution of Midterm-1

Salil Koner

October 16, 2019

Problem 1

(a) Let $\mathbf{Z}_i = (Z_{i1}, Z_{i2}, Z_{i3}) \sim IID\text{Multinomial}\,(1, p_1, p_2, p_3)$ $\quad p_1+p_2+p_3 = 1$ independent from $X_{ij} \sim f_j(x;\theta)$ independent across i, j.

Then, Y_i defined by

$$Y_i = Z_{i1}X_{i1} + Z_{i2}X_{i2} + Z_{i3}X_{i3}$$

is distributed as a mixture of the 3 components specified by the problem.

(b) The complete data likelihood comes from the contribution of both $\mathbf{Y}$ and $\mathbf{Z}$, which is given by

$$\begin{aligned} L(\boldsymbol{\theta}, p_1, p_2, p_3) &= \Pi_{i=1}^n f_{(Y_i, \mathbf{Z}_i)}(y_i, \mathbf{z}_i; \boldsymbol{\theta}, \mathbf{p}) \\ &= \Pi_{i=1}^n f_{Y_i|\mathbf{Z}_i}(y_i; z_i; \boldsymbol{\theta}) f_{\mathbf{Z}_i}(\mathbf{z}_i; \mathbf{p}) \\ &= \Pi_{i=1}^n \left[f_1(y_i|\boldsymbol{\theta})^{Z_{i1}} f_2(y_i|\boldsymbol{\theta})^{Z_{i2}} f_3(y_i|\boldsymbol{\theta})^{1-Z_{i1}-Z_{i2}}\right] \left[p_1^{Z_{i1}} p_2^{Z_{i1}} (1-p_1-p_2)^{1-Z_{i1}-Z_{i2}}\right] \end{aligned}$$

So, the complete data log-likelihood is

$$l_c(\boldsymbol{\theta}, p_1, p_2, p_3) = \sum_{i=1}^n \sum_{j=1}^2 Z_{ij}\left[\log f_j(y_i|\boldsymbol{\theta}) + \log p_j\right] + \sum_{i=1}^n (1 - Z_{i1} - Z_{i2}) \log\left[f_3(y_i|\boldsymbol{\theta}) + \log(1-p_1-p_2)\right]$$

(c)

$$\begin{aligned} Q(\boldsymbol{\theta}, \mathbf{p}|\mathbf{Y}, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}) &= E\; l_c(\boldsymbol{\theta}, p_1, p_2, p_3) \,|\mathbf{Y}, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}} \\ &= \sum_{i=1}^n \sum_{j=1}^2 \mathrm{E}\left(Z_{ij}|Y_i, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}\right)\left[\log f_j(y_i|\boldsymbol{\theta}) + \log p_j\right] \\ &\quad + \sum_{i=1}^n \left(1 - \mathrm{E}\left(Z_{i1}|Y_i, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}\right) - \mathrm{E}\left(Z_{i2}|Y_i, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}\right)\right) \log\left[f_3(y_i|\boldsymbol{\theta}) + \log(1-p_1-p_2)\right] \\ &= \sum_{i=1}^n \sum_{j=1}^2 w_{ij}^{\nu}\left[\log f_j(y_i|\boldsymbol{\theta}) + \log p_j\right] + \sum_{i=1}^n (1 - w_{i1}^{\nu} - w_{i2}^{\nu}) \log\left[f_3(y_i|\boldsymbol{\theta}) + \log(1-p_1-p_2)\right] \end{aligned}$$

where

$$\begin{aligned} w_{ij}^{\nu} &= \mathrm{E}\left(Z_{ij}|Y_i, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}\right) = P\left(Z_{ij} = 1|Y_i, \mathbf{p}^{\boldsymbol{\nu}}, \boldsymbol{\theta}^{\boldsymbol{\nu}}\right) \\ &= \frac{f_j(Y_i|\boldsymbol{\theta}^{\nu}) p_j^{\nu}}{\sum_{j=1}^3 f_j(Y_i|\boldsymbol{\theta}^{\nu}) p_j^{\nu}} \end{aligned}$$

Problem 2

(a) The score function is

$$S(\boldsymbol{\theta}) = b(\mathbf{y}) - 2c\boldsymbol{\theta}$$

The Fisher information matrix is

$$I(\boldsymbol{\theta}) = -\mathrm{E}\left(S'(\boldsymbol{\theta})\right) = 2c\,\mathbf{I}_b > 0 \quad \text{as } c > 0$$

(b) MLE of the solution of the score function, this means

$$\boldsymbol{\theta}_{\text{MLE}} = \frac{1}{2c} b(\mathbf{y})$$

and the maximizer is confirmed by the fact that

$$S'(\boldsymbol{\theta}) = -2c\,\mathbf{I}_b \prec 0 \quad \text{as } c > 0$$

Wald test is

$$T_w = 2c\left(\widehat{\boldsymbol{\theta}}_{\text{MLE}} - \boldsymbol{\theta}_0\right)^\top (\boldsymbol{\theta}_{\text{MLE}} - \boldsymbol{\theta}_0)$$

(c) The score test is

$$T_s = \frac{1}{2c}\left(b(\mathbf{y}) - 2c\boldsymbol{\theta}_0\right)^\top \left(b(\mathbf{y}) - 2c\boldsymbol{\theta}_0\right)$$

(d)

$$\ell(\boldsymbol{\theta}_0) = a(\mathbf{y}) + b(\mathbf{y})^\top \boldsymbol{\theta}_0 - c\boldsymbol{\theta}_0^\top \boldsymbol{\theta}_0$$

and

$$\ell\left(\widehat{\boldsymbol{\theta}}_{\text{MLE}}\right) = a(\mathbf{y}) + \frac{1}{2c} b(\mathbf{y})^\top b(\mathbf{y})$$

The LRT is given by

$$T_{\text{LR}} = -2\left[b(\mathbf{y})^\top \boldsymbol{\theta}_0 - c\boldsymbol{\theta}_0^\top \boldsymbol{\theta}_0 - \frac{1}{2c} b(\mathbf{y})^\top b(\mathbf{y})\right] = T_s + \frac{1}{2c} b(\mathbf{y})^\top b(\mathbf{y})$$

(e) The null distribution of all the statistics are asymptotically same

$$T_w, T_s, T_{\text{LR}} \sim \chi_b^2 \quad \text{asymptotically under } H_0$$

Problem 3

(a) The likelihood function is

$$L(\theta_1, \theta_2) = (\theta_1 + \theta_2)^{-n} \Pi_{i=1}^n \left[\exp(-y_i/\theta_1)\right]^{1(y_i > 0)} \left[\exp(y_i/\theta_2)\right]^{1(y_i \le 0)}$$

(b) The log-likelihood function is

$$\ell(\theta_1, \theta_2) = -n\log(\theta_1 + \theta_2) - \frac{z_1}{\theta_1} + \frac{z_2}{\theta_2}$$

(c) the score function is

$$S(\theta_1, \theta_2) = \left(-\frac{n}{\theta_1 + \theta_2} + \frac{z_1}{\theta_1^2}\ ,\quad -\frac{n}{\theta_1 + \theta_2} - \frac{z_2}{\theta_2^2}\right)^\top$$

(d) The MLE is the solution of score equation, this implies

$$\frac{z_1}{\theta_1^2} = \frac{n}{\theta_1 + \theta_2} = -\frac{z_2}{\theta_2^2} \implies \theta_1 = c\theta_2 \quad \text{where } c = \sqrt{-\frac{z_1}{z_2}}$$

Putting that in one of the equation we get,

$$\frac{z_1}{c^2\theta_2^2} = \frac{n}{c\theta_2 + \theta_2} \implies \hat{\theta}_2 = \frac{(c+1)z_1}{nc^2} \ , \ \hat{\theta}_1 = \frac{(c+1)z_1}{nc}$$

putting the value of c we get

$$\hat{\theta}_1 = \left(1 + \sqrt{-\frac{z_2}{z_1}}\right)\frac{z_1}{n} \quad \hat{\theta}_2 = \left(1 + \sqrt{-\frac{z_1}{z_2}}\right)\frac{z_2}{n}$$

(e) To calculate the Fisher information matrix we need to calculated $E(Z_1)$ and $E(Z_2)$.

$$\begin{aligned}
E\left(Y\ 1(Y>0)\right) &= \frac{1}{\theta_1+\theta_2}\int_0^\infty y\exp(-y/\theta_1)dy \\
&= \frac{\theta_1}{\theta_1+\theta_2}\int_0^\infty \frac{y}{\theta_1}\exp(-y/\theta_1)dy \\
&= \frac{\theta_1}{\theta_1+\theta_2}E(W) \qquad [\text{where } W \sim \exp(\theta_1)] \\
&= \frac{\theta_1^2}{\theta_1+\theta_2}
\end{aligned}$$

Similarly,

$$\begin{aligned}
E\left(Y\ 1(Y\le 0)\right) &= \frac{1}{\theta_1+\theta_2}\int_{-\infty}^0 y\exp(y/\theta_2)dy \\
&= \frac{\theta_2}{\theta_1+\theta_2}\int_{-\infty}^0 \frac{y}{\theta_2}\exp(y/\theta_2)dy \\
&= \frac{\theta_2}{\theta_1+\theta_2}E(-W) \qquad [\text{where } W \sim \exp(\theta_2)] \\
&= -\frac{\theta_2^2}{\theta_1+\theta_2}
\end{aligned}$$

Because Y_i's are iid

$$E(Z_1) = \frac{n\theta_1^2}{\theta_1+\theta_2} \qquad E(Z_2) = -\frac{n\theta_2^2}{\theta_1+\theta_2}$$

$$\frac{\partial^2 \ell}{\partial(\theta_1,\theta_2)} = \begin{pmatrix} \frac{n}{(\theta_1+\theta_2)^2} - \frac{2z_1}{\theta_1^3} & \frac{n}{(\theta_1+\theta_2)^2} \\ \frac{n}{(\theta_1+\theta_2)^2} & \frac{n}{(\theta_1+\theta_2)^2} + \frac{2z_2}{\theta_2^3} \end{pmatrix}$$

This implies,

$$I(\theta_1,\theta_2) = -\frac{n}{(\theta_1+\theta_2)^2}\mathbf{J}_2 + \frac{2n}{(\theta_1+\theta_2)}\operatorname{diag}\left(\theta_1^{-1},\theta_2^{-1}\right)$$

(f) By property of MLE (as all the regularity conditions hold true)

$$\sqrt{n}\left[\begin{pmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{pmatrix} - \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}\right] \longrightarrow \mathrm{N}_2\left(\mathbf{0}, I(\theta_1,\theta_2)^{-1}\right)$$

Problem 4

Since $X_n = \mathcal{O}_p(n)$ it implies $X_n = nZ_{1n}$, where $Z_{1n} = O_p(1)$. Similarly $Y_n = o_p(n)$ implies $Y_n = nZ_{2n}$, where $Z_{2n} \to_p 0$.

(a) Answer $O_p(n)$. Intuition: $X_n + Y_n = n(Z_{1n} + Z_{2n})$, where the term $Z_{1n} + Z_{2n}$ is also contained in a compact interval for almost all values of n with a high probability. This means $X_n + Y_n = n\mathcal{O}_p(1) = \mathcal{O}_p(n)$.

(b) Answer $o_p(n^2)$. Intuition: $X_nY_n = n^2Z_{1n}Z_{2n}$, and $Z_{1n}Z_{2n} \to_p 0$.

Note: In part (b), because, convergence in probability implies bounded in probability, X_nY_n is also $\mathcal{O}_p(n^2)$, but it is not the best answer given the amount of information provided.